

AD-A158 542

IDENTIFIABILITY UNDER APPROXIMATION FOR AN ELLIPTIC
BOUNDARY VALUE PROBLEM. (U) BROWN UNIV PROVIDENCE RI
LEFSCHETZ CENTER FOR DYNAMICAL SYSTE.

1/1

UNCLASSIFIED

K KUNISCH ET AL. APR 85 LCDS-85-13

F/G 12/1

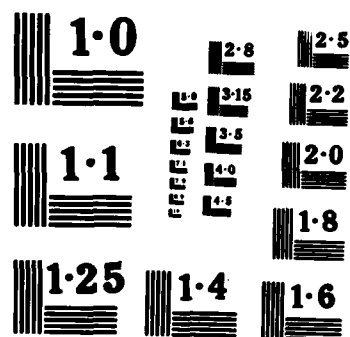
NL



END

FILMED

DTIC



NATIONAL BUREAU OF STANDARDS
MICROCOPY RESOLUTION TEST CHART



Lefschetz Center for Dynamical Systems

Approved for publ
distribution unl

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE

AD A158542

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION Unclassified		1b. RESTRICTIVE MARKINGS	
2a. SECURITY CLASSIFICATION AUTHORITY		3. DISTRIBUTION AVAILABILITY OF REPORT Approved for public release; distribution unlimited	
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE			
4. PERFORMING ORGANIZATION REPORT NUMBER		5. MONITORING ORGANIZATION REPORT NUMBER AFOSR-TR-85-0554	
6a. NAME OF PERFORMING ORGANIZATION Leifsonetz Center for Dynamical Systems	6b. OFFICE SYMBOL If applicable	7a. NAME OF MONITORING ORGANIZATION AFOSR	
6c. ADDRESS (City, State and ZIP Code) Division of Applied Mathematics Brown University, Providence, RI 02910		7b. ADDRESS (City, State and ZIP Code) Bldg. 410 Bolling AFB, D.C. 20332-6448	
8a. NAME OF FUNDING SPONSORING ORGANIZATION AFOSR	8b. OFFICE SYMBOL If applicable NM	9. ACQUISITION INSTRUMENT IDENTIFICATION NUMBER AFOSR-84-0298	
9c. ADDRESS (City, State and ZIP Code) Bldg. 410 Bolling AFB, D.C. 20332-6448		10. SOURCE OF FUNDING NOS.	
		PROGRAM ELEMENT NO. 61102F	PROJECT NO. 2304
		TASK NO. A1	WORK UNIT NO.
11. TITLE (Include Security Classification) Identifiability Under Approximation for an Elliptic Boundary Value Problem			
12. PERSONAL AUTHOR(S) Karl Kunisch and L. W. White			
13a. TYPE OF REPORT Interim	13b. TIME COVERED FROM TO	14. DATE OF REPORT (Yr., Mo., Day) April 1985	15. PAGE COUNT 36
16. SUPPLEMENTARY NOTATION			
17. COSATI CODES		18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)	
FIELD	GROUP	SUB GR	
XXXXXX	XXXXXXXXXX	XXXX	
		Galerkin approximations, output least squares	
19. ABSTRACT (Continue on reverse if necessary and identify by block number) Necessary and sufficient conditions for identifiability of the diffusion coefficient in Galerkin approximations to a two point boundary value problem are derived for various subspaces of Galerkin sub-spaces. The results are further used to investigate output least squares identifiability and output least squares stability of the diffusion coefficient.			
20. DISTRIBUTION STATEMENT (For AFOSR only)		21. ABSTRACT SECURITY CLASSIFICATION	
UNCLASSIFIED UNLIMITED (X) SAME AS APT (X) OTHERS ()		UNCLASSIFIED	
22a. NAME OF RESPONSIBLE INDIVIDUAL Dr. Marc D. Jacobus		22b. TELEPHONE NUMBER (Include Area Code) 301-767-1241	22c. OFFICE SYMBOL NM

IDENTIFIABILITY UNDER APPROXIMATION FOR AN
ELLIPTIC BOUNDARY VALUE PROBLEM

by

Karl Kunisch and L. W. White

April 1985

LCDS #85-13

DTIC
ELECTE
AUG 29 1985
S D G

DISTRIBUTION STATEMENT A

Approved for public release;
Distribution Unlimited

IDENTIFIABILITY UNDER APPROXIMATION FOR AN
ELLIPTIC BOUNDARY VALUE PROBLEM

by

Karl Kunisch*

Division of Applied Mathematics
Brown University
Providence, Rhode Island

and

Institut für Mathematik
Technische Universität Graz
A-8010 GRAZ, AUSTRIA

and

L. W. White**

Department of Mathematics
The University of Oklahoma
Norman, Oklahoma 73019

April 1985

Accession For	
NTIS GRA&I	☒
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A/1	



* Research in part supported by the Fonds zur Förderung der Wissenschaftlichen Forschung, Austria, No. S3206, and by Air Force Office of Scientific Research under contract No. AF-AFOSR 84-0398 (5-28393).

** Research in part supported by the Air Force Office of Scientific Research under contract No. AF-AFOSR 84-0271.

A

Abstract

Necessary and sufficient conditions for identifiability of the diffusion coefficient in Galerkin approximations to a two point boundary value problem are derived for various choices of Galerkin subspaces. The results are further used to investigate output least squares identifiability and output least squares stability of the diffusion coefficient.

Key Words: Parameter estimation; stability of output least squares problem, boundary value problem.

Analysis; matrices (mathematics)

Identifiability Under Approximation for an Elliptic Boundary Value Problem

1. Introduction

In this note we study the following boundary value problems:

$$\begin{aligned} &-(au_x)_x + cu = f \quad \text{in } (0,1) \\ (1.1) \quad &u_x(0) - u_x(1) = 0. \end{aligned}$$

Let $I = (0,1)$ and $f \in L^2(I)$. Recall that if $a \in H^1(I)$ with $a(x) \geq \alpha > 0$ and $c \in L^2(I)$ with $c(x) \geq \underline{c} > 0$ a.e., then there exists a unique solution $u = u(a)$ of (1.1) in $H^2(I)$. We are concerned with the identification of the coefficient a , given information of the solution $u(a)$ and in particular we will study the injectivity of the mapping $a^M \rightarrow u^N(a^M)$ where a^M is some approximation to a and u^N an approximation to u . At first we describe the problem in a more general context.

Let $\mathcal{L}: H^2(I) \rightarrow Z$ be a continuous linear operator from the solution space to the observation space Z describing the type of available information of the state u . To determine the coefficient corresponding to an observation $z \in Z$ of the system that is modelled by (1.1) the following output least squares formulation is used frequently:

$$(1.2) \quad \underset{Q_{ad}}{\text{minimize}} \quad |\mathcal{L}u(a) - z|_Z^2.$$

Here the set Q_{ad} of admissible parameters is chosen such that the existence of a solution of (1.2) is guaranteed.

For example, if n observations $\{z_k\}_{k=1}^n$ taken at the points $\{x_k\}_{k=1}^n$ are available, we may take $Z = \mathbb{R}^n$ with $\mathcal{L}: H^2(I) \rightarrow \mathbb{R}^n$ defined by $\mathcal{L}u = \{u(x_k)\}_{k=1}^n$. In this case (1.2) becomes

$$(1.3) \quad \underset{Q_{ad}}{\text{minimize}} \sum_{k=1}^n |u(x_k; a) - z_k|^2.$$

Alternatively one might have distributed observations $z \in L^2(I)$, or using the data $\{z_k\}_k^n$ at $\{x_k\}_{k=1}^n$ one might want to obtain a function $z \in L^2(I)$ either by interpolation or least squares regression; for $\mathcal{L}$ we would then take $\mathcal{L}u = u$ and the optimization problem becomes

$$(1.4) \quad \underset{Q_{ad}}{\text{minimize}} |u(a) - z|_{L^2}^2.$$

In either case an appropriate choice for Q_{ad} would be

$$Q_{ad} = \{a \in H^1: a(x) \geq \alpha > 0, \quad |a|_{H^1} \leq \delta\},$$

with $\gamma > \alpha$, [6].

Defining the attainable set $\mathcal{V} = \{\mathcal{L}u(a): a \in Q_{ad}\}$ one may view the optimization problem (1.2) as having two parts:

- (i) given $z \in Z$ find z_{proj} , the projection of z on $\mathcal{V}$,
- (ii) given z_{proj} find $\bar{a} \in Q_{ad}$ such that $\mathcal{L}u(\bar{a}) = z_{proj}$.

Assuming the existence of z_{proj} , the uniqueness of z_{proj} depends on the geometry of $\mathcal{V}$. In (ii) there exists an $\bar{a}$ such that $\mathcal{L}u(\bar{a}) = z_{proj}$ by definition of Q_{ad} . The question of uniqueness of such an $\bar{a}$ arises and it is guaranteed if $\phi: a \rightarrow \mathcal{L}u(a)$ is injective at $\bar{a}$.

Injectivity of Φ at $\bar{a}$ is called identifiability of a at $\bar{a}$. The above mentioned uniqueness problems are rather involved in general, see, e.g., [2; 7, Appendix], and [11] for a hyperbolic equation.

When solving (1.2) on a computer it is necessary to replace (1.1)-(1.2) by a finite dimensional problem. This is done by approximating both the solutions of (1.1) and the set Q_{ad} by functions from finite dimensional function spaces. A finite dimensional version of the minimization problem (1.2) is then solved to obtain an estimate for the unknown coefficient a (compare e.g., [1], [5]). Again the existence and uniqueness questions analogous to the two steps (i) and (ii) above can be considered.

The main purpose of this investigation is the study of the uniqueness for the finite dimensional analog of (ii). If for a chosen approximation of a by a^M the mappings $a^M \rightarrow u^N(a^M)$ is injective at $\bar{a}^M$, then a is called identifiable under approximation at $\bar{a}^M$. The related question for parabolic equations in dimension one has been treated in [4].

Our results below indicate that the injectivity of a^M depends upon certain rank conditions that imply compatibility conditions upon the spaces used to approximate the coefficient a and the solution $u(a)$. It will be seen that a may be identifiable under approximation without the known sufficient conditions for identifiability of a in (1.1) being satisfied [9]. The results here, although depending on the choice of Neumann boundary conditions can easily be adapted to different boundary conditions.

In section 2 we formulate the discrete problems and give general conditions for identifiability under approximation. In section 3 we examine several concrete examples and obtain necessary and sufficient conditions

for identifiability under approximation for these cases. Identifiability will be guaranteed if there is a sufficient amount of movement in the observations. On the other hand, if the coefficient is assumed to be known at points where the observations are stationary, then it can still be identifiable at the remaining parts of the domain $(0,1)$. Section 4 is devoted to the problem of continuous dependence of the solution of the discretized version of (1.3) or (1.4) on the observation z and Q_{ad} . Sufficient conditions for output least squares identifiability (OLSI) [2] and output least squares stability (OLSS) [3] are given. Finally, in section 5 we report the findings of a numerical experiment that supports the practical relevance of our results.

2. Basic Results

To approximate (1.2) by the standard finite element method [10] let $\{B_i\}_{i=0}^N$ and $\{\phi_j\}_{j=1}^M$ be sets of linearly independent functions defined on I with $B_i \in H^1(I)$ and ϕ_j piecewise continuous. Let $A^M = \text{span}\{\phi_j: j = 1, \dots, M\}$ and $H^N = \text{span}\{B_i: i = 0, \dots, N\}$. Setting

$$\bar{u}^N = \sum_{i=0}^N \mu_i B_i \quad \text{and} \quad a^M = \sum_{j=1}^M a_j^M \phi_j$$

we have upon integration by parts of (1.1) with u replaced by $\bar{u}^N$:

$$\sum_{i=0}^N \mu_i \langle a B_{i,x}, B_{k,x} \rangle + \sum_{i=0}^N \mu_i \langle c B_i, B_k \rangle = \langle f, B_k \rangle \quad \text{for } k = 0, \dots, N,$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product in L^2 . Replacing a by a^M it follows that

$$\sum_{i=0}^N \mu_i \sum_{j=1}^M a_j^M \langle \phi_j B_{i,x}, B_{k,x} \rangle + \sum_{i=0}^N \mu_i \langle c B_i, B_k \rangle = \langle f, B_k \rangle,$$

for $k = 0, \dots, N$.

Rearranging the summations in this last expression we arrive at

$$(2.1) \quad \sum_{j=1}^M a_j \sum_{i=0}^N \langle \phi_j B_{i,x}, B_{k,x} \rangle \mu_i + \sum_{i=0}^N \langle c B_i, B_k \rangle \mu_i = \langle f, B_k \rangle$$

for $k = 0, \dots, N$.

We now make the following definitions: H_j and K are $(N+1) \times (N+1)$ matrices with the (i,k) -th elements given by

$$(H_j)_{i,k} = \langle \phi_j B_{i,x}, B_{k,x} \rangle \quad \text{and} \quad (K)_{i,k} = \langle c B_i, B_k \rangle,$$

for $i, k = 0, \dots, N$; $j = 1, \dots, M$. Similarly $\tilde{f} \in \mathbb{R}^{N+1}$, $\tilde{u} \in \mathbb{R}^{N+1}$ and $\tilde{a}^M \in \mathbb{R}^M$ are given by

$$(\tilde{f})_k = \langle f, B_k \rangle, \quad (\tilde{u})_i = u_i \quad \text{and} \quad (\tilde{a}^M)_j = a_j^M.$$

With this notation (2.1) becomes

$$(2.2) \quad \sum_{j=1}^M a_j^M H_j \tilde{u} + K \tilde{u} = \tilde{f},$$

where we used the symmetry of H_j and K . Thus we obtain a mapping $\tilde{a}^M \rightarrow \tilde{u}(\tilde{a}^M)$ from $\mathbb{R}^M$ into $\mathbb{R}^{N+1}$ or, equivalently, $a^M \rightarrow u^N(a^M)$,

$a^M = \sum_{j=1}^M a_j^M \phi_j$, from A^M to H^N , that is well defined as long as $\sum_{j=1}^M a_j^M H_j + K$ is invertible. For example, if $c \geq 0$ and $\{B_{i,x}\}_{i=0}^N$ are linearly independent, then $\sum_{j=1}^M a_j^M H_j + K$ is invertible for all

$$\tilde{a}^M \in A = \{\tilde{a}^M \in \mathbb{R}^M: \sum_{j=1}^M (\tilde{a}^M)_j \phi_j > 0 \text{ on } I\}.$$

Similarly, if $c \geq \gamma > 0$ as assumed throughout, then again $\sum_j a_j H_j + K$ is invertible for all $\tilde{a}^M \in A$.

We now define identifiability of $\tilde{a}^M = \text{col}(a_1^M, \dots, a_k^M)$ in (2.2).

Definition 2.1. The parameter $\tilde{a}^M \in A$ in (2.2) is called identifiable if $\tilde{\mu}(\tilde{a}^M) = \tilde{\mu}(\tilde{b}^M)$ implies $\tilde{a}^M = \tilde{b}^M$ for all $\tilde{b}^M \in A$.

For a specific choice of approximation of a in (1.1) by a^M , we say that a is identifiable under approximation at a^M if a^M is identifiable.

Theorem 2.1. Let $\{\phi_j\}_{j=0}^M$ and $\{B_i\}_{i=0}^N$ be linearly independent, $c(x) \geq \underline{c} > 0$ and $\tilde{a}^M \in A$. Then $\tilde{a}^M \in A$ is identifiable if and only if the vectors $\{H_j \tilde{\mu}(\tilde{a}^M)\}_{j=1}^M$ are linearly independent.

Proof. Using (2.2), linear independence of $H_j \tilde{\mu}(\tilde{a}^M)$ clearly implies identifiability of $\tilde{a}^M$. Conversely assume that there exists a nontrivial vector $(\alpha_1, \dots, \alpha_M) \in \mathbb{R}^M$ with

$$(2.3) \quad \sum_{j=1}^M \alpha_j H_j \tilde{\mu}(\tilde{a}^M) = 0.$$

Then $\sum_{j=1}^M (a_j^M - \epsilon \alpha_j) \phi_j > 0$ for some sufficiently small $\epsilon > 0$ and $\tilde{a}_1^M$ given by $(\tilde{a}_1^M)_j = a_j^M - \epsilon \alpha_j$ satisfies $\tilde{a}_1^M \in A$. Multiplying (2.3) by ϵ and subtracting it from (2.2) we find that $\tilde{\mu}(\tilde{a}^M) = \tilde{\mu}(\tilde{a}_1^M)$. This ends the proof.

Since $H_j: \mathbb{R}^{N+1} \rightarrow \mathbb{R}^{N+1}$, for $j = 1, \dots, M$, we have the following:

Corollary 2.1. If $M > N+1$, then $\tilde{a}^M$ in (2.2) is not identifiable.

Corollary 2.2. If $\tilde{u}(\tilde{a}^M) \in \text{Ker}(H_j)$ for some $j = 1, \dots, M$, then $\tilde{a}^M$ is not identifiable.

Proof. If $\tilde{u}(\tilde{a}^M) \in \text{Ker}(H_j)$ then the set $\{H_j \tilde{u}(\tilde{a}^M)\}_{j=1}^M$ is linearly dependent and the result follows from Theorem 2.1.

Remark 2.1. Corollary 2.2 should be compared with the condition

$|u_x| \geq k_1 > 0$ which is known to be a sufficient condition for identifiability of a in the infinite dimensional problem (1.1) [9].

Definition 2.2. The coordinates $\{\tilde{a}_{j_k}^M\}_{k=1}^{\hat{M}}$, $\hat{M} \leq M$, of the parameter vector $\tilde{a}^M \in A$ are called identifiable if $\tilde{u}(\tilde{a}^M) = \tilde{u}(\tilde{b}^M)$ and $\tilde{a}_j^M = \tilde{b}_j^M$ for all $j \neq j_k$, $k = 1, \dots, \hat{M}$ imply $\tilde{a}^M = \tilde{b}^M$, for all $\tilde{b}^M \in A$.

Proposition 2.1. The coordinates $\{\tilde{a}_{j_k}^M\}_{k=1}^{\hat{M}}$ of $\tilde{a}^M \in A$ are identifiable if and only if the vectors $\{H_{j_k} \tilde{u}(\tilde{a}^M)\}_{k=1}^{\hat{M}}$ are linearly independent.

The proof is obvious from that of Theorem 2.1.

3. Several Examples

In this section we consider several concrete examples and determine their identifiability properties. We point out that here we use N to denote the number of subintervals of I and N and M of the previous section are a function of this N .

Case 1. Let I be partitioned into N subintervals of length $1/N$.

For $i = 0, \dots, N$ define the linear spline basis functions

$$(3.1) \quad B_i(x) = \begin{cases} Nx-i+1, & \frac{i-1}{N} \leq x \leq \frac{i}{N}, \\ -Nx+i+1, & \frac{i}{N} \leq x \leq \frac{i+1}{N}, \\ 0 & \text{otherwise,} \end{cases}$$

and for $j = 1, \dots, N$ the 0-th order splines

$$(3.2) \quad \phi_j(x) = \begin{cases} 1, & \frac{j-1}{N} \leq x \leq \frac{j}{N}, \\ 0 & \text{otherwise.} \end{cases}$$

Thus N and M of the previous section are both N here.

We approximate the solution u of (1.1) by linear splines and the coefficient a by constant splines. It is straightforward to compute the $(N+1) \times (N+1)$ matrices H_j , $j = 1, \dots, N$:

$$H_j = N \begin{pmatrix} 0 & \dots & \dots & \overset{j-1}{\vdots} & \dots & \dots & \overset{j}{\vdots} & \dots & \dots & 0 \\ \vdots & \ddots & 0 & & & 0 & & & & \vdots \\ \vdots & & & 1 & -1 & & & & & \vdots \\ \vdots & & & -1 & 1 & & & & & \vdots \\ \vdots & & & & & 0 & & & & \vdots \\ 0 & \dots & \dots & & & 0 & \dots & \dots & 0 \end{pmatrix} \begin{matrix} \dots j-1 \\ \dots j \end{matrix}.$$

Let $\tilde{\mu} = \tilde{\mu}(\tilde{a}^M) = \text{col}(\mu_0, \dots, \mu_N)$ with $\tilde{a}^M \in A$. Then

$$H_j \tilde{\mu} = N \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \mu_{j-1} - \mu_j & \dots j-1 \\ -\mu_{j-1} + \mu_j & \dots j \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Now set $\beta_j = \mu_j - \mu_{j-1}$ for $j = 1, \dots, N$. To study the linear independence of the vectors $\{H_j \tilde{\mu}\}_{j=1}^N$, note that

$$(H_1 \tilde{\mu}, \dots, H_N \tilde{\mu}) = NB$$

where B is the $(N+1) \times N$ matrix

$$B = \begin{pmatrix} -\beta_1 & 0 & \cdots & 0 \\ \beta_1 & -\beta_2 & & \\ 0 & \beta_2 & & \\ \vdots & & \ddots & \\ 0 & & & -\beta_{N-1} \\ & & & \beta_{N-1} & -\beta_N \\ & & & 0 & \beta_N \end{pmatrix}$$

Lemma 3.1. The vectors $\{H_j \tilde{u}\}_{j=1}^N$ are linearly independent if and only if $\beta_i \neq 0$ for all $i = 1, \dots, N$.

Proof. It is easily shown that B is row equivalent [8] to

$$\tilde{B} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & & \\ \vdots & & \ddots & \\ 0 & & & 1 \\ 0 & & & 0 \end{pmatrix},$$

provided $\beta_i \neq 0$ for all i and $\{H_j \tilde{u}\}_{j=1}^N$ are linearly independent in this case. Conversely, if $\beta_i = 0$ for some i , then the column rank of B is less than N and linear dependence of $\{H_j \tilde{u}\}_{j=1}^N$ follows.

Theorem 3.1. In Case 1, $\tilde{a}^M$ is identifiable, if and only if $\mu_i(\tilde{a}^M) \neq \mu_{i-1}(\tilde{a}^M)$ for all $i = 1, \dots, N$.

Proposition 3.1. The coordinates $\{\tilde{a}_{j_k}^M\}_{k=1}^{\hat{M}}$ of $\tilde{a}^M$ are identifiable if and only if $\mu_{j_k}(\tilde{a}^M) \neq \mu_{j_k-1}(\tilde{a}^M)$ for all $k = 1, \dots, \hat{M}$.

The interpretation of this result is that the parameter $\tilde{a}^M$ can only be identified at coordinates where the corresponding observations are non-stationary.

Case 2. Let N be even and let I be partitioned into subintervals of length $\frac{1}{N}$. The functions B_i , $i = 0, \dots, N$, are taken as in (3.1). Here, however, we define the functions ϕ_j for $j = 1, \dots, \frac{N}{2}$ by

$$\phi_j(x) = \begin{cases} 1, & \frac{2(j-1)}{N} \leq x \leq \frac{2j}{N}, \\ 0 & \text{otherwise.} \end{cases}$$

Thus M of section 2 is $\frac{N}{2}$ now. We find in this case that the $(N+1) \times (N+1)$ matrices H_j , for $j = 1, \dots, \frac{N}{2}$ are given by

$$H_j = N \begin{pmatrix} 0 & \vdots & & \vdots & & 0 \\ \vdots & & & & & \vdots \\ \vdots & & 1 & -1 & 0 & \vdots \\ \vdots & & -1 & 2 & -1 & \vdots \\ \vdots & & 0 & -1 & 1 & \vdots \\ 0 & \vdots & & \vdots & & 0 \end{pmatrix},$$

where the first entry of the nontrivial submatrix is in the $2j-2, 2j-2$ position of H_j . With $\tilde{\mu} = \tilde{\mu}(\tilde{a}^M) = \text{col}(\mu_0, \dots, \mu_N)$ as before we have

$$H_j \tilde{\mu} = N \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \mu_{2j-2} - \mu_{2j-1} \\ -\mu_{2j-2} + 2\mu_{2j-1} - \mu_{2j} \\ -\mu_{2j-1} + \mu_{2j} \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

where the first nonzero entry occurs in the $2j-2$ coordinate. Setting

$\beta_i = \mu_i - \mu_{i-1}$ for $i = 1, \dots, N$ we thus have

$$H_j \tilde{\mu} = N \begin{pmatrix} 0 \\ \vdots \\ 0 \\ -\beta_{2j-1} \\ \beta_{2j-1} - \beta_{2j} \\ \beta_{2j} \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

for $j = 1, \dots, \frac{N}{2}$. To investigate the linear independence of $\{H_j \tilde{\mu}\}_{j=1}^{N/2}$, note that

$$(H_1 \tilde{\mu}, \dots, H_{N/2} \tilde{\mu}) = NB$$

where the $(N+1) \times \frac{N}{2}$ -matrix B is given by

$$B = \begin{pmatrix} -\beta_1 & 0 & 0 \\ \beta_1 - \beta_2 & 0 & 0 \\ \beta_2 & -\beta_3 & 0 \\ 0 & \beta_3 - \beta_4 & \dots \\ 0 & \beta_4 & 0 \\ 0 & 0 & \beta_{N-1} \\ 0 & 0 & \beta_{N-1} - \beta_N \\ 0 & 0 & \beta_N \end{pmatrix},$$

and the collection of vectors $\{H_j \tilde{u}\}_{j=1}^{N/2}$ is linearly independent if and only if $\text{rank}(B) = \frac{N}{2}$.

Lemma 3.2. In the Case 2 the vectors $\{H_j \tilde{u}\}_{j=1}^{N/2}$ are linearly independent if and only if $\beta_{2i-1} \neq 0$ or $\beta_{2i} \neq 0$ for $i = 1, \dots, \frac{N}{2}$.

Proof. The matrix B is row equivalent to

$$\tilde{B} = \begin{pmatrix} \beta_1 & 0 & 0 \\ \beta_2 & 0 & 0 \\ 0 & \beta_3 & 0 \\ | & \beta_4 \dots & | \\ 0 & 0 & 0 \\ | & | & \beta_{N-1} \\ 0 & 0 & \beta_N \end{pmatrix}.$$

From this and the fact that the dimension of the column space of a matrix is equal to the rank of that matrix, the result follows.

Theorem 3.2. In Case 2, $\tilde{a}^M$ is identifiable if and only if $\mu_{2i-1} \neq \mu_{2i-2}$ or $\mu_{2i} \neq \mu_{2i-1}$ for $i = 1, \dots, \frac{N}{2}$.

Case 3. Let I be partitioned into N subintervals of length $\frac{1}{N}$.

Again we take the functions B_i , $i = 0, \dots, N$ to be those defined in (3.1). Further we set $\phi_i = B_i$, $i = 0, \dots, N$. Thus M of section 2 is $N+1$ here and both ϕ_j and B_i are linear splines defined on the same mesh. In this case the structure of the $(N+1) \times (N+1)$ matrices $\{H_j\}_{j=0}^N$

is slightly more complicated than in Cases 1 and 2. These matrices are now given as follows:

$$H_0 = \frac{N}{2} \begin{pmatrix} 1 & -1 & \vdots & & \\ -1 & 1 & \vdots & & \\ \vdots & \vdots & \ddots & & \\ & & & \bigcirc & \\ & & & & \bigcirc \end{pmatrix},$$

with the first entry in the (0,0)-element. For $j = 1, \dots, N-1$ we have

$$H_j = \frac{N}{2} \begin{pmatrix} \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix},$$

where the first entry of the nontrivial submatrix appears in the (j-1, j-1) position of H_j . Finally

$$H_N = \frac{N}{2} \begin{pmatrix} \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix},$$

where the first nonzero entry occurs in the N-1, N-1 element. The vectors $\{H_j \tilde{u}\}_{j=0}^N$ are given by

$$H_0 \tilde{\mu} = \frac{N}{2} \begin{pmatrix} -(\mu_1 - \mu_0) \\ \mu_1 - \mu_0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad H_N \tilde{\mu} = \frac{N}{2} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ -(\mu_N - \mu_{N-1}) \\ \mu_N - \mu_{N-1} \end{pmatrix}$$

$$H_j \tilde{\mu} = \frac{N}{2} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \mu_{j-1} - \mu_j \\ -\mu_{j-1} + 2\mu_j - \mu_{j+1} \\ -\mu_j + \mu_{j+1} \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

where for $j = 1, \dots, N-1$ the first entry occurs in row $j-1$. Setting

$\beta_i = \mu_i - \mu_{i-1}$ for $i = 1, \dots, N$ we have

$$H_0 \tilde{\mu} = \frac{N}{2} \begin{pmatrix} -\beta_1 \\ \beta_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad H_N \tilde{\mu} = \frac{N}{2} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ -\beta_N \\ \beta_N \end{pmatrix}, \quad H_j \tilde{\mu} = \frac{N}{2} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ -\beta_j \\ \beta_j - \beta_{j+1} \\ \beta_{j+1} \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

for $j = 1, \dots, N-1$. To investigate the linear independence of $\{H_j \tilde{u}\}_{j=0}^N$ note that

$$(H_0 \tilde{u}, \dots, H_N \tilde{u}) = \frac{N}{2} B,$$

where the $(N+1) \times (N+1)$ matrix B is given by

$$B = \begin{pmatrix} -\beta_1 & -\beta_1 & 0 & 0 \\ \beta_1 & \beta_1 - \beta_2 & 0 & 0 \\ 0 & \beta_2 & 0 & 0 \\ \vdots & 0 & \vdots & \vdots \\ 0 & 0 & \dots & 0 \\ & & & -\beta_{N-1} & 0 \\ & & & \beta_{N-1} - \beta_N & -\beta_N \\ & & & \beta_N & \beta_N \end{pmatrix}.$$

Performing row operations on B we find the B is equivalent to

$$\tilde{B} = \begin{pmatrix} \beta_1 & \beta_1 & 0 & 0 & 0 \\ 0 & \beta_2 & \beta_2 & 0 & 0 \\ \vdots & 0 & \beta_3 & \vdots & \vdots \\ & & 0 & \dots & 0 \\ & & & & \beta_{N-1} & 0 \\ & & & & \beta_N & \beta_N \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

We then see that B has rank less or equal to N and therefore we have

Theorem 3.3. In case 3, $\tilde{a}^M \in A$ is not identifiable.

If one decreases the number of ϕ_j 's in Case 3 then it is reasonable to expect that sufficient and necessary conditions for the identifiability of $\tilde{a}^M$ in the spirit of Cases 1 and 2 can be obtained. We verify this next for a particular choice of N and M . Moreover, in Section 5 we present a numerical experiment which tends to support this contention.

Case 4. Let I be partitioned into $2N$ subintervals of length $1/2N$.

We choose the functions B_i for $i = 0, \dots, 2N$ as

$$B_i(x) = \begin{cases} 2Nx - i + 1, & \frac{i-1}{2N} \leq x \leq \frac{i}{2N}, \\ -2Nx + i + 1, & \frac{i}{2N} \leq x \leq \frac{i+1}{2N}, \\ 0 & \text{otherwise.} \end{cases}$$

For the functions ϕ_j we take

$$\phi_j(x) = \begin{cases} Nx - j + 1, & \frac{j-1}{N} \leq x \leq \frac{j}{N}, \\ -Nx + j + 1, & \frac{j}{N} \leq x \leq \frac{j+1}{N}, \\ 0 & \text{otherwise,} \end{cases}$$

for $j = 0, \dots, N$; thus M of section 2 is $N+1$. The $(2N+1) \times (2N+1)$ -matrices $\{H_j\}_{j=0}^N$ are given as follows:

$$H_0 = \frac{N}{2} \begin{pmatrix} 3 & -3 & 0 & & & \\ -3 & 4 & -1 & & \circ & \\ 0 & -1 & 1 & & & \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ & \circ & & \vdots & \circ & \\ & & & \vdots & & \ddots \end{pmatrix},$$

with 3 in the (0,0) element,

$$H_j = \frac{N}{2} \begin{pmatrix} & & & \circ & & \\ & & & & & \vdots \\ \vdots & & & & & \\ & & 1 & -1 & 0 & 0 & 0 \\ & & -1 & 4 & -3 & 0 & 0 \\ \circ & & 0 & -3 & 6 & -3 & 0 \\ & & 0 & 0 & -3 & 4 & -1 \\ & & 0 & 0 & 0 & -1 & 1 \\ & & & & & & \vdots \\ & & & \circ & & & \vdots \end{pmatrix},$$

where the first entry of the nontrivial submatrix occurs in the $2(i-1)$, $2(i-1)$ -element and

$$H_N = \frac{N}{2} \begin{pmatrix} & & & \vdots & & \\ & \circ & & \vdots & \circ & \\ & & & \vdots & & \\ \vdots & & & \vdots & & \\ & & & \vdots & & \\ & \circ & & 1 & -1 & 0 \\ & & & -1 & 4 & -3 \\ & & & 0 & -3 & 3 \end{pmatrix}.$$

Let $\tilde{\mu} = \tilde{\mu}(\tilde{a}^M) = \text{col}(\mu_0, \dots, \mu_{2N})$. Then we see that

$$H_0 \tilde{u} = \frac{N}{2} \begin{pmatrix} 3\mu_0 - 3\mu_1 \\ -3\mu_0 + 4\mu_1 - \mu_2 \\ -\mu_1 + \mu_2 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad H_N \tilde{u} = \frac{N}{2} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \mu_{2N-2} - \mu_{2N-1} \\ -\mu_{2N-2} + 4\mu_{2N-1} - 3\mu_{2N} \\ -3\mu_{2N-1} + 3\mu_{2N} \end{pmatrix}$$

and for $j = 1, \dots, N-1$

$$H_j \tilde{u} = \frac{N}{2} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \mu_{2i-2} - \mu_{2i-1} \\ -\mu_{2i-2} + 4\mu_{2i-1} - 3\mu_{2i} \\ -3\mu_{2i-1} + 6\mu_{2i} - 3\mu_{2i+1} \\ -3\mu_{2i} + 4\mu_{2i+1} - \mu_{2i+2} \\ -\mu_{2i+1} + \mu_{2i+2} \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

where the first nonzero entry occurs in row $2(i-1)$. To investigate the linear independence of $\{H_j \tilde{u}\}_{j=0}^N$ we put $\beta_i = \mu_i - \mu_{i-1}$ for $i = 1, \dots, 2N$ and observe that

$$(H_0 \tilde{u}, \dots, H_N \tilde{u}) = \frac{N}{2} B$$

where the $(2N+1) \times (N+1)$ matrix B is given by

$$B = \begin{pmatrix} -3\beta_1 & -\beta_1 & 0 & & & \\ 3\beta_1 - \beta_2 & \beta_1 - 3\beta_2 & 0 & & & \\ \beta_2 & 3\beta_2 - 3\beta_3 & -\beta_3 & & & \\ 0 & 3\beta_3 - \beta_4 & \beta_3 - 3\beta_3 & & & \\ & \beta_4 & 3\beta_4 - 3\beta_5 & & & \\ & 0 & 3\beta_5 - \beta_6 & \dots & & \\ & & \beta_6 & & 0 & \\ & & 0 & & -\beta_{2N-3} & \\ & & & & \beta_{2N-3} - 3\beta_{2N-2} & 0 \\ & & & & 3\beta_{2N-2} - 3\beta_{2N-1} & -\beta_{2N-1} \\ & & & & 3\beta_{2N-1} - \beta_{2N} & \beta_{2N-1} - 3\beta_{2N} \\ & & & & \beta_{2N} & 3\beta_{2N} \end{pmatrix}.$$

Performing row operations on B we obtain an equivalent matrix $\tilde{B}$:

$$\tilde{B} = \begin{pmatrix} 3\beta_1 & \beta_1 & 0 & & 0 & 0 \\ \beta_2 & 3\beta_2 & 0 & & & \\ 0 & 3\beta_3 & \beta_3 & & & \\ & \beta_4 & 3\beta_4 & & & \\ & 0 & 3\beta_5 & & & \\ & & \beta_6 & \dots & & \\ & & 0 & & \beta_{2N-3} & 0 \\ & & & & 3\beta_{2N-2} & 0 \\ & & & & 3\beta_{2N-1} & \beta_{2N-1} \\ & & & & \beta_{2N} & 3\beta_{2N} \end{pmatrix},$$

which is of dimension $2N \times (N+1)$. The vectors $H_j \tilde{u}$ are linearly independent if $\text{rank}(\tilde{B}) = N+1$.

Theorem 3.4. In Case 4, if $\mu_{2i-1} \neq \mu_{2i-2}$ or $\mu_{2i} \neq \mu_{2i-1}$ for all $i = 1, \dots, N$ and $(\mu_{2i-1} - \mu_{2i-2})(\mu_{2i} - \mu_{2i-1}) \neq 0$ for some $i = 1, \dots, N$, where $\tilde{u} = \tilde{u}(\tilde{a}^M)$, then $\tilde{a}^M \in A$ is identifiable.

Proof. Let $\tilde{\alpha} = \text{col}(\tilde{\alpha}_0, \dots, \tilde{\alpha}_N) \in \mathbb{R}^{N+1}$ and $\tilde{B}\tilde{\alpha} = 0$. Choose i_1 such that $\beta_{2i_1-1} \cdot \beta_{2i_1} \neq 0$. Then $\alpha_{i_1-1} = \alpha_{i_1} = 0$. Further $\alpha_i = 0$ for all other i , since $\beta_{2i-1} \neq 0$ or $\beta_{2i} \neq 0$. This implies linear independence of the columns of $\tilde{B}$ and thus of B . Theorem 3.1 then implies the result.

Theorem 3.5. In Case 4, if $\mu_1 - \mu_0 = \mu_2 - \mu_1 = 0$ or $\mu_{N-1} - \mu_{N-2} = \mu_N - \mu_{N-1} = 0$ or $\mu_{2i-1} - \mu_{2i-2} = \mu_{2i} - \mu_{2i-1} = \mu_{2i+1} - \mu_{2i} = \mu_{2i+2} - \mu_{2i+1} = 0$ for some $i = 2, \dots, N-2$, where $\tilde{u} = \tilde{u}(\tilde{a})$ then $\tilde{a}^M \in A$ is not identifiable.

Proof. Under the assumptions of the theorem the column-rank of $\tilde{B}$ is not maximal and this implies the result.

Remark 3.1. Four consecutive zeros in the β_i 's do not necessarily imply nonidentifiability, provided the zeros start with an even index and are not at the "beginning" or "end" of the sequence $\{\beta_i\}$; in particular $\beta_{2i} = \beta_{2i+1} = \beta_{2i+2} = \beta_{2i+3} = 0$ with $2 \leq i < N-2$ does not imply nonidentifiability. For example let $N = 5$, $\beta_1 = \beta_2 = \beta_3 = \beta_8 = \beta_9 = \beta_{10} = 1$ and $\beta_4 = \beta_5 = \beta_6 = \beta_7 = 0$. Then $\{H_j \tilde{u}\}$, $j = 0, \dots, N$, with $\beta_i = \mu_i - \mu_{i-1}$, are linearly independent. Choosing μ_0 and $\tilde{a}^M$, we can thus calculate μ_i , $i = 1, \dots, 2N$ and $\tilde{f}$ such that $\tilde{a}^M$ is identifiable and $\beta_{2i} = \dots = \beta_{2i+3} = 0$.

Remark 3.2. The conditions of Theorems 3.1 - 3.4 are conditions on the variation of adjacent $\mu_i(\tilde{a}^M)$ -values. If this variation is sufficient, the identifiability of $\tilde{a}^M$ is guaranteed. The results indicate that the larger the difference between the dimension of the state space approximation and the dimension of the parameter space approximation is, the more likely it is that identifiability of the approximated coefficient holds. In [9] identifiability of a in (1.1) is studied under various conditions on the sign of u_x and u_{xx} . The most general condition implying identifiability of a is $\inf_I (\max |u_x|, u_{xx}) > 0$. Clearly one can construct examples where this condition is not met but identifiability under approximation, e.g., according to one of the cases 1 - 4, of $\tilde{a}^M$ holds.

4. Two Stability Concepts

In this section we discuss the application of two concepts of stability to the finite dimensional output least squares problem

$$(P_M^N) \quad \text{minimize } \|u^N(a^M) - z\|_2^2 \text{ over } C,$$

where C is a convex and closed subset of

$$Q_{ad}^M = \{a^M = \sum_{j=1}^M a_j^M \phi_j : a_j^M \in \mathbb{R}, a^M(x) \geq \alpha > 0, \|a^M\|_{H^1} \leq \gamma\},$$

and $u^N(a^M) = \sum_{i=0}^N \nu_i B_i$, with $\text{col}(\nu_0, \dots, \nu_N) = \tilde{u}$ satisfying (2.2). The existence of a minimum a_*^M of (P_M^N) can easily be argued. The reason for introducing the set $C \subset Q_{ad}$ here is that for the first stability concept, uniqueness of solutions of (P_M^N) is required and this cannot be guaranteed over all of Q_{ad} . Here we consider H^1 -smooth approximating coefficients, some remarks on L^∞ approximations are given further below. We investigate the continuous dependence of a_*^M on z and also on Q_{ad}^M when dealing with the second stability concept. For $C \subset Q_{ad}^M$ let $\mathcal{V}(C) = \{u^N(a^M) : a^M \in C\}$ denote the attainable set.

Definition 4.1. [2] The parameter a^M in (2.2) is called output least squares identifiable (OLSI) by (P_M^N) over $C \subset Q_{ad}^M$, if there exists a neighborhood $\tilde{\mathcal{V}}$ of $\mathcal{V}(C)$ such that for every $z \in \tilde{\mathcal{V}}$ the problem (P_M^N) has a unique solution a_*^M depending continuously on z .

Let $A_{inj} \subset Q_{ad}$ be such that $\{H_j \tilde{u}(\tilde{a}^M)\}$ is linearly independent for every $a^M = \sum_{j=1}^M (\tilde{a}^M)_j \phi_j \in A_{inj}$. As examples for such sets we can

take neighborhoods in Q_{ad} of points of identifiability in the sense of section 2.

Theorem 4.1. Let A_{inj} be as just described. Then a^M in (2.2) is OLSI by (P_M^N) over every closed convex subset C of A_{inj} , provided that $\text{diam } C$ is sufficiently small and z is sufficiently close to the $\mathcal{V}(C)$. More precisely, (P_M^N) has a unique solution a_*^M depending Lipschitz-continuously on z as long as $\text{dist}(z, \mathcal{V}(C))$ is sufficiently small.

Proof. By ([2], Theorem 4) it suffices to show that $a^M \rightarrow u^N(a^M)$ is twice continuously Fréchet differentiable with $a^M \rightarrow u_{a^M}^N(a^M)$ injective on C . This is equivalent to the existence of continuous first and second order derivatives of $\tilde{a}^M \rightarrow \tilde{u}^N(\tilde{a}^M)$ with $\tilde{a}^M$ such that

$\sum_{j=1}^M (\tilde{a}^M)_j \phi_j = a^M \in A_{inj}$ and injectivity of $\tilde{u}_{\tilde{a}^M}^N(\tilde{a}^M)$ for every $a^M \in C$.

Let $\tilde{u}_{\tilde{a}^M}^N(\tilde{a}^M; \tilde{h}) = \tilde{\eta}$ and $\tilde{u}_{\tilde{a}^M}^N(\tilde{a}^M; \tilde{h}, \tilde{k}) = \tilde{\xi}$ be the first, resp. second, derivative in directions $\tilde{h}$ and $(\tilde{h}, \tilde{k})$. Then

$$(4.1) \quad \tilde{\eta} = -L^{-1}(\Sigma h_j \quad H_j \tilde{u}^M(\tilde{a}^M))$$

and

$$(4.2) \quad \tilde{\xi} = -L^{-1}(\Sigma k_j \quad H_j \tilde{\eta} + \Sigma h_j H_j \tilde{u}_{\tilde{a}^M}^N(\tilde{a}^M; \tilde{k})),$$

where $L = \Sigma a_j^M H_j + K$ and $\tilde{h} = \text{col}(h_1, \dots, h_M)$, $\tilde{k} = \text{col}(k_1, \dots, k_M)$,

and the continuity assumptions follow. The injectivity of $\tilde{u}_{\tilde{a}^M}^N(\tilde{a}^M)$ is guaranteed by linear independence of $\{H_j \tilde{u}^M(\tilde{a}^M)\}$; this ends the proof.

To describe the second notion of stability we consider the case

$$C = Q_{ad}^M:$$

$$(P_M^N)_w \quad \text{minimize} \quad |u^N(a^M) - z|^2 \quad \text{over} \quad Q_{ad}^M.$$

We study continuous dependence of local solutions of $(P_M^N)_w$ on $w = (z, \alpha, \delta) \in W$, where $W = H^0 \times \mathbb{R} \times \mathbb{R}$. Here W is endowed with the Hilbert-space product norm. We always assume $0 < \alpha < \gamma$, so that Q_{ad}^M is not empty and solutions of (2.2) and $(P_M^N)_w$ exist.

Definition 4.2. [2] The parameter a^M is called output least squares (OLS)-stable in Q_{ad}^M at the local solution a_0^M of $(P_M^N)_{w^0}$, $w^0 \in W$, if there exists a neighborhood $V(w^0)$ of w^0 in W , a neighborhood $V(a_0^M)$ of a_0^M in H^1 and a constant κ , such that for all $w = (z, \alpha, \gamma) \in V(w^0)$ there exists a local solution a_w^M of $(P_M^N)_w$ with $a_w^M \in V(a_0^M)$ and for all local solutions $a_w^M \in V(a_0^M)$ of $(P_M^N)_w$ we have

$$|a_w^M - a_0^M|_{H^1} \leq \kappa |w - w^0|_W^{1/2}.$$

Remark 4.1. In comparing OLSI to OLS-stability we observe the following differences: OLSI requires uniqueness of the solutions of the minimization problem, whereas for OLS-stability, uniqueness is not required, with continuity being checked at each local solution. If OLSI holds, then the solutions depend on the observations in a Lipschitz continuous way, whereas OLS-stability only guarantees Hölder continuous dependence. Further, OLSI requires continuous dependence of the solutions on the observation only, whereas OLS-stability involves continuous dependence on the observations as well as on the admissible set Q_{ad} .

Output least squares stability is proved by techniques that guarantee stability of solutions of abstract optimization problems with respect to perturbations in the problem data, see [2] and the references given there. Let $A^M = \text{span} \{\phi_j\}_{j=1}^M$ and let $F(a^M)$ be the Lagrange functional associated with (P_M^N) :

$$F(a^M) = |u^N(a^M) - z|^2 - \lambda^* g(a^M),$$

where $\lambda^* \in C^* \times \mathbb{R}$ and

$g: A^M \times W \rightarrow C^* \times \mathbb{R}$ is given by

$$g(a^M, w) = (\alpha - a^M, |a^M|_{H^1}^2 - \gamma^2), \quad w = (z, \alpha, \gamma).$$

Note that $a^M \in Q_{ad}(w)$ if and only if $g(a^M, w) \in \hat{K} = C_- \times \mathbb{R}_-$, with C_- and $\mathbb{R}_-$ the natural negative cones in $C(I)$ and $\mathbb{R}$. We shall frequently drop the index w and write $g(a^M)$ and Q_{ad}^M for $g(a^M, w)$ and $Q_{ad}^M(w)$.

Theorem 4.2. Let $A^M = \text{span}\{\phi_j\}_{j=1}^M$ be such that it contains the constant functions, let $(z^0, \alpha^0, \gamma^0) = w^0 \in W$ with $0 < \alpha^0 < \gamma^0$ and let $a_0^M = \sum_{j=1}^M (\tilde{a}_0^M)_j \phi_j$ be a local solution of $(P_M^N)_{w^0}$. If $\{H_j \tilde{u}(\tilde{a}_0^M)\}_{j=1}^M$ are linearly independent vectors in $\mathbb{R}^{N+1}$ and $|u^N(a_0^M) - z|$ is sufficiently small, then a^M is OLS-stable in $Q_{ad}^M(w^0)$ at the local solution a_0^M of $(P_M^N)_{w^0}$.

For the proof of this theorem the following lemma on the regularity of the constraint set Q_{ad}^M will be required; its proof is quite similar to that of Lemma 4.2 in [2] but will be included for the sake of completeness.

Lemma 4.1. Let A^M contain the constant functions. Then every $a^M \in Q_{ad}^M$ is a regular point, i.e., $0 \in \text{int}\{g(a^M) + \mathcal{R}(g_{a^M}(a^M)) - \hat{K}\} \subset C \times \mathbb{R}$, where $\mathcal{R}$ denotes the range of the mapping $g_{a^M}(a^M)$.

Proof of Lemma 4.1. We need to show that

$$(4.3) \quad \begin{aligned} 0 &\in \text{int}\{g(a^M) + g_{a^M}(a^M)A^M - C_- \times \mathbb{R}_-\} \\ &= \text{int}\{\alpha - a^M - h^M + C_+, |a^M|_{H^1}^2 - \gamma^2 + 2\langle a^M, h^M \rangle_{H^1} + \mathbb{R}_+ : h^M \in A^M\}, \end{aligned}$$

where we used that $g_{a^M}(a^M)h^M = (-h^M, 2\langle a^M, h^M \rangle_{H^1})$. Let $(\phi, r) \in C \times \mathbb{R}$ with $|(\phi, r)|_{C \times \mathbb{R}} < \delta$ and $\delta > 0$ to be chosen sufficiently small. Note that $\phi - \min \phi \in C_+$, and $\min \phi \in A^M$. In view of the first component in (4.3) we decompose ϕ as

$$\phi = \alpha - a^M - (\alpha - a^M - \min \phi) + \phi - \min \phi$$

and therefore $\phi \in \alpha - a^M - A^M + C_+$. As for the second component in (4.3) observe that

$$\begin{aligned} |a^M|_{H^1}^2 - \gamma^2 + 2\langle a^M, \alpha - a^M - \min \phi \rangle_{H^1} &= -|a^M|_{H^1}^2 + 2\langle a^M, \alpha - \min \phi \rangle_{H^1} - \gamma^2 \\ &\leq \alpha^2 - \gamma^2 + 2\delta |a^M|_{H^1}. \end{aligned}$$

Thus, for δ sufficiently small one can choose $\tilde{r} \in \mathbb{R}_+$ such that

$$r = |a^M|_{H^1}^2 - \gamma^2 + 2\langle a^M, \alpha - a^M - \min \phi \rangle_{H^1} + \tilde{r}$$

and, since (ϕ, r) was arbitrary, a^M is shown to be a regular point.

Proof of Theorem 4.2. We apply results on the stability of abstract optimization problems as summarized in section 3 of [2]. Due to the fact that $a^M \rightarrow |u^N(a^M) - z^0|^2$ and $a^M \rightarrow g(a^M, w^0)$ are twice continuously at a_0^M and since the point $a_0^M \in Q_{ad}^M$ is a regular point, it suffices to establish a lower bound on the second derivative of F at $\bar{a}^M$. Let $\eta = u_{M,a}^N(a_0^M; h^M)$ and $\xi = u_{M,a}^N(a_0^M; h^M, h^M)$ for $h^M \in A^M$. Then

$$\begin{aligned} F_{a,a}^{M,M}(a_0^M; h^M, h^M) &= \langle u^N(a_0^M) - z, \xi \rangle_{H^0} + |\eta|_{H^0}^2 - 2\lambda |h^M|_{H^1}^2 \\ &\geq -|u^N(a_0^M) - z|_{H^0} |\xi|_{H^0} + |\eta|_{H^0}^2 - 2\lambda |h^M|_{H^1}^2, \end{aligned}$$

where $\lambda \leq 0$ is the Lagrange multiplier associated with the norm constraint. In view of (4.1), (4.2), the finite dimensionality of A^M , and the linear independence of $\{H_j \tilde{u}(\tilde{a}_0^M)\}_{j=1}^M$ it follows that there exist constants c_1 and c_2 such that

$$F_{a,a}^{M,M}(a_0^M; h^M, h^M) \geq -c_1 |u^N(a_0^M) - z|_{H^0} |h^M|_{H^1}^2 + c_2 |h^M|_{H^1}^2,$$

so that for $|u^N(a_0^M) - z|$ sufficiently small there exists a constant c_3 with

$$F_{a,a}^{M,M}(a_0^M; h^M, h^M) \geq c_3 |h^M|_{H^1}^2;$$

from which the result follows [2; Theorem 3.2, 3.3].

Remark 4.2. If $Q_{ad}^M \subset L^\infty$ only and $|a^M|_{H^1} \leq \gamma$ is replaced by

$|a^M|_{L^\infty} \leq \gamma$ in the definition of Q_{ad}^M , then again one can show existence of solutions of (P_M^N) and Theorem 4.1 holds with obvious modifications.

The results leading to Theorem 4.2 need yet to be generalized to handle the nondifferentiable L^∞ -norm constraint.

5. Numerical Results

In this section we present some results of a numerical experiment to estimate the coefficient a in (1.1) given observations z of u . To solve (1.2) with $C = I$ we consider (2.2) which defines a mapping $\tilde{a}^M \rightarrow \tilde{u}(\tilde{a}^M)$ for $\tilde{a}^M \in A$, and the finite dimensional minimization problems

$$(5.1) \quad \text{minimize} \quad \int_0^1 \left(\sum_{i=0}^N u_i(\tilde{a}^M) B_i - z \right)^2 dx.$$

For our experiments we imposed no constraints on $\tilde{a}^M \in \mathbb{R}^M$, although $\tilde{u}(\tilde{a}^M)$ is not well defined for some $\tilde{a}^M$. The basis functions ϕ_j and B_i were chosen as linear spline functions with equidistant grid on $(0,1)$. As data z we took the values of a solution of (1.1) by choosing the coefficient a and the observation $z(x) = u(x) = x^2(1-x)^2$, and calculating $f = u - (au_x)_x$ from it. Using this f we then compute $\tilde{u}$ from (2.2) as we solve (5.1). For the minimization the Newton-Raphson algorithm was used.

In our calculations $N = 10$ represents the number of subintervals used in the linear spline approximation for the solution of (1.1). Thus the dimension of the approximation space for the solution is 11. Further NBI is the number of subintervals of I that determine the linear spline approximation of a ; the dimension of the approximation space for a is $NBI + 1$. A necessary condition for identifiability of $\tilde{a}^M$ is thus $NBI + 1 \leq 11$, see Corollary 2.1. We show calculations for $NBI = 4, 5, 6, 8 - 11$, for the choice of $a(x) = 1 + x$. In the first five cases good results are obtained. Note that $NBI = 5$ and $N = 10$ is a special

case of Theorem 3.4. In the case $NBI = 10$, $\tilde{a}^M$ is not identifiable by Theorem 3.3. Numerically this is reflected by the appearance of oscillations as NBI approaches 10 from below, see the graphs for $NBI = 9, 10, 11$. The start-up value for the minimization routine was chosen as $\tilde{a}_0^M \equiv 2$. We point out that a different scaling of the axes in the various graphs was utilized. We also show the graphs for $u^N(a^M)$, when $N = 10$ and $NBI = 9, 10, 11$. The -graphs for the approximating solutions for $NBI = 4, 5, 6, 8$ are indistinguishable from $NBI = 9$.

FIGURE 1

1D FLOW - NEUMAN B.C.S
 $A=1+X$, $U=X \cdot X \cdot (1.0-X)^{**2}$
 $N=10$, $NBI=4$

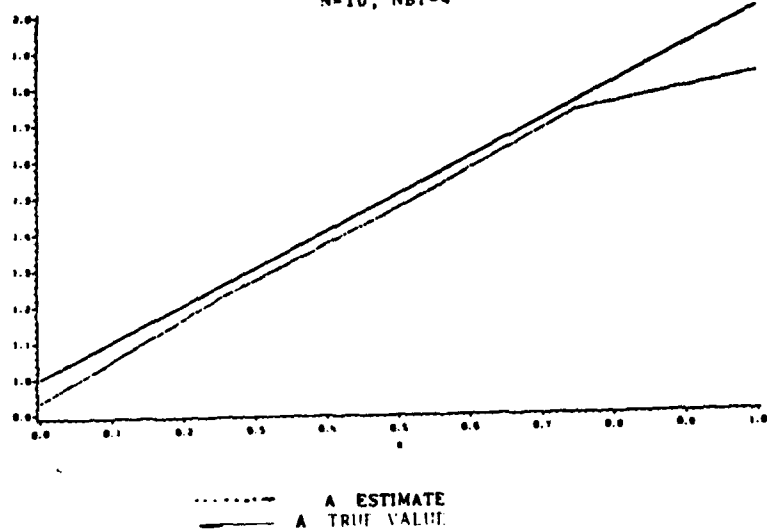


FIGURE 2

1D FLOW - NEUMAN B.C.S
 $A=1+X$, $U=X \cdot X \cdot (1.0-X)^{**2}$
 $N=10$, $NBI=5$

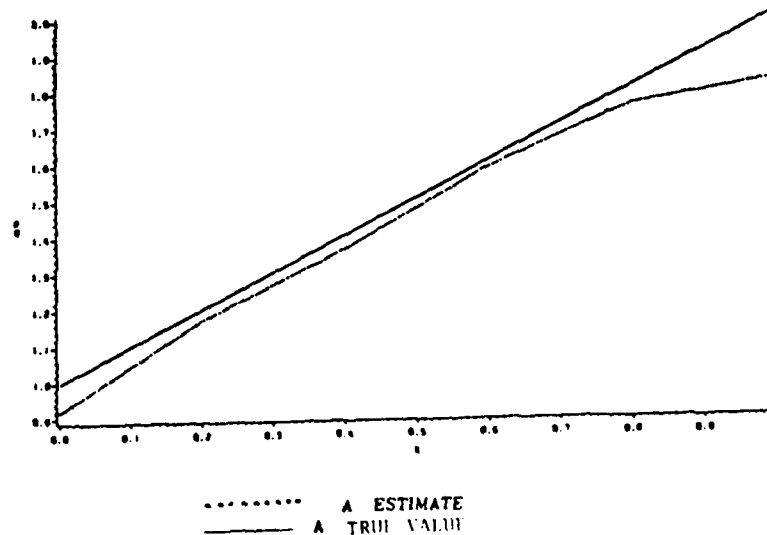


FIGURE 3

1D FLOW - NEUMAN B.C.S
 $A=1+X$, $U=X \cdot X \cdot (1.0-X)^{**2}$
 $N=10$, $NBI=6$

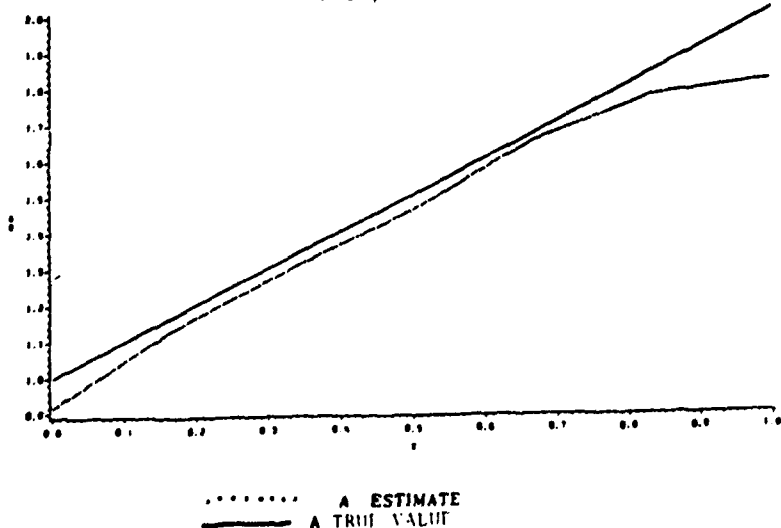


FIGURE 4

1D FLOW - NEUMAN B.C.S
 $A=1+X$, $U=X \cdot X \cdot (1.0-X)^{**2}$
 $N=10$, $NBI=8$

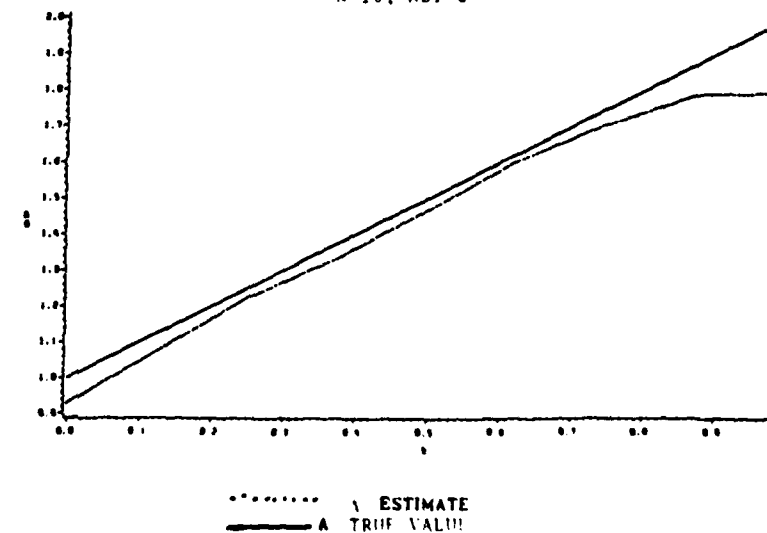
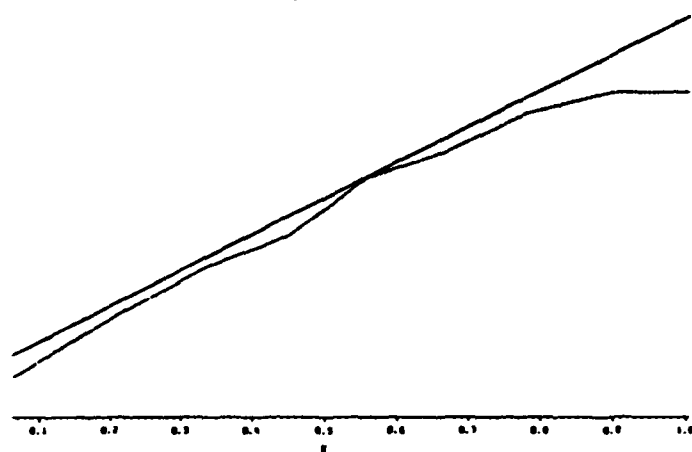


FIGURE 5

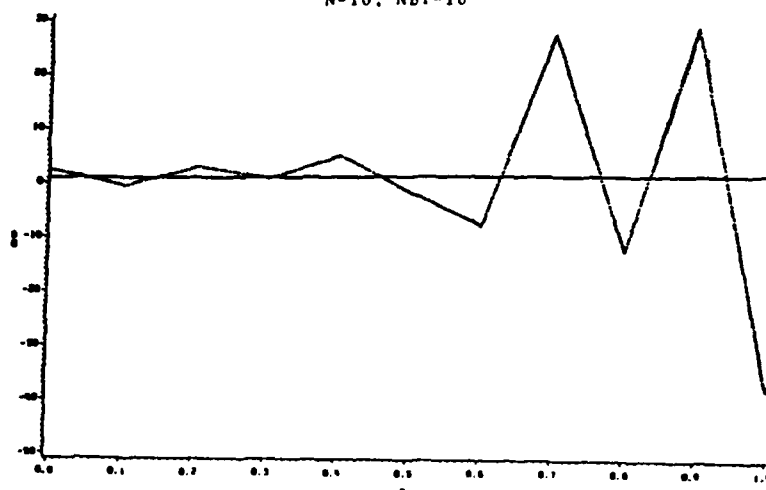
1D FLOW - NEUMAN B.C.S
 $A=1-X$, $U=X \cdot X \cdot (1.0-X)^{**2}$
 $N=10$, $NBI=9$



..... A ESTIMATE
 ——— A TRUE VALUE

FIGURE 6

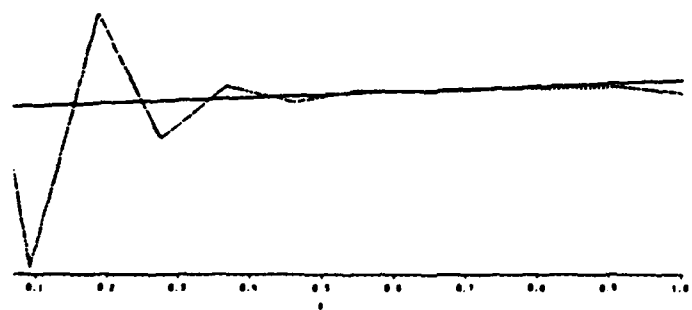
1D FLOW - NEUMAN B.C.S
 $A=1-X$, $U=X \cdot X \cdot (1.0-X)^{**2}$
 $N=10$, $NBI=10$



..... A ESTIMATE
 ——— A TRUE VALUE

FIGURE 7

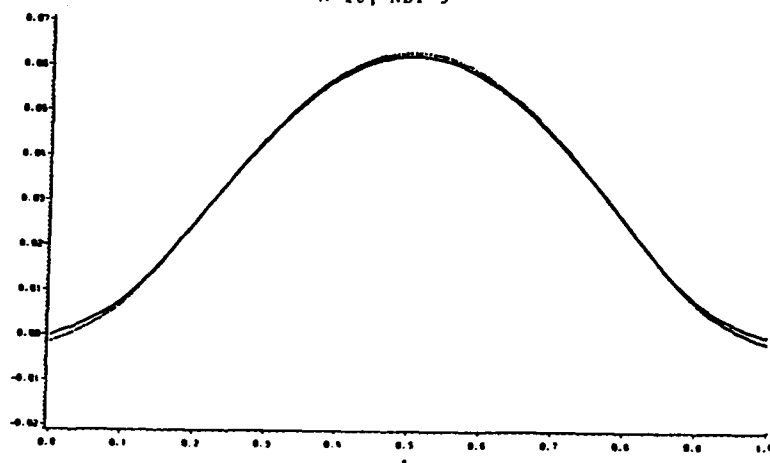
1D FLOW - NEUMAN B.C.S
 $A=1-X$, $U=X \cdot X \cdot (1.0-X)^{**2}$
 $N=10$, $NBI=11$



..... A ESTIMATE
 ——— A TRUE VALUE

FIGURE 8

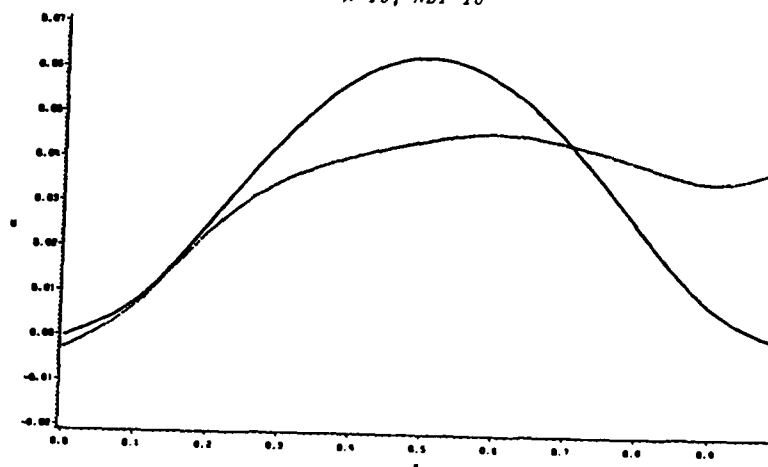
1D FLOW - NEUMAN B.C.S
 $A=1-X$, $U=X \cdot X \cdot (1-X)^{**2}$
 $N=10$, $NBI=9$



..... U ESTIMATE
 _____ U TRUE VALUE

FIGURE 9

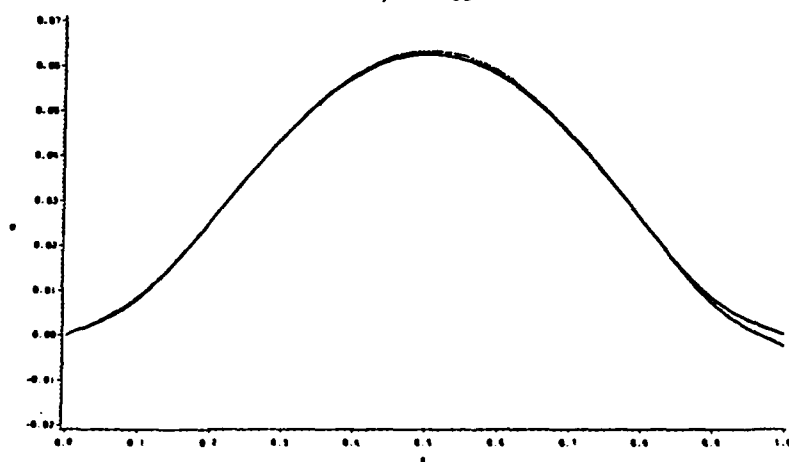
1D FLOW - NEUMAN B.C.S
 $A=1-X$, $U=X \cdot X \cdot (1-X)^{**2}$
 $N=10$, $NBI=10$



..... U ESTIMATE
 _____ U TRUE VALUE

FIGURE 10

1D FLOW - NEUMAN B.C.S
 $A=1-X$, $U=X \cdot X \cdot (1-X)^{**2}$
 $N=10$, $NBI=11$



..... U ESTIMATE
 _____ U TRUE VALUE

References

- [1] Banks, H. T., P. L. Daniel and E. S. Armstrong: Parameter estimation for a static model of Maypole-hoop column antenna surfaces, ICASE Report 82-26, 1982.
- [2] Chavent, G.: Local stability of the output least square parameter estimation technique, INRIA Research Report No. 136.
- [3] Colonius, F. and K. Kunisch: Stability for parameter estimation in two point boundary value problems, submitted.
- [4] Kunisch, K. and L. W. White: Parameter Identifiability Under approximation, to appear in Quarterly of Appl. Math.
- [5] Kunisch, K. and L. W. White: Parameter estimation for elliptic equations in multi-dimensional domains with point and flux observations, to appear in Nonlinear Analysis, Theory, Methods and Applications.
- [6] Kunisch, K. and L. W. White: Regularity properties in parameter estimation of diffusion coefficients in one dimensional elliptic boundary value problems, submitted.
- [7] Levitan, B. J.: Distribution of Zeroes of Entire Functions, AMS Translations 5, Providence, 1964.
- [8] Noble, B.: Applied Linear Algebra, Prentice-Hall Inc., Englewood Cliffs, 1969.
- [9] Richter, G. R.: An inverse problem for the steady state diffusion equation, SIAM J. Appl. Math., 41 (1981), 210-221.
- [10] Schultz, M.: Spline Analysis, Prentice Hall, Englewood Cliffs, 1973.
- [11] White, L. W.: Identification of a Friction Coefficient in a First Order Linear Hyperbolic Equation, Proceedings of the 1983 IEEE Control and Decision Conference.

END

FILMED

10-85

DTIC